Supporting Information

"Click" synthesis of small molecule-peptide conjugates for organellespecific delivery and inhibition of lysosomal cysteine proteases

Yuhui Loh,^a Haibin Shi,^a Mingyu Hu,^a Shao Q. Yao*^{a,b,c}

Departments of Chemistry^a and Biological Sciences,^b NUS MedChem Program of the Life Sciences Institute,^c National University of Singapore, Singapore 117543

Contents

- 1. General information
- 2. Synthetic procedure and LC-MS profile of alkyne-functionalized inhibitor
- 3. Synthetic procedure and LC-MS profiles of azide-functionalized localization peptides
- 4. Synthetic procedure and LC-MS profiles of "click" inhibitor-peptide conjugates
- 5. Procedure for IC_{50} evaluation against recombinant cathepsin L
- 6. Cell culture
- 7. Procedure and results for western blot of HepG2 cell lysate
- 8. Procedure for IC₅₀ evaluation against HepG2 cell lysate
- 9. Procedure and results for cell imaging with control peptides
- 10. Procedure and results for cell imaging with cell-permeable fluorogenic substrate
- 11. Procedure and results for flow cytometry analysis
- 12. ¹H and ¹³C NMR spectra

1. General information

All ¹H NMR spectra were taken on a Bruker ACF-300/500 MHz NMR spectrometer, using CDCl₃ or (CD₃)₂SO as the solvent. Chemical shifts are reported in parts per million referenced with respect to residual solvent (CDCl₃ = 7.26 ppm and (CD₃)₂SO = 2.50 ppm). LC-MS spectra were recorded using Shimadzu LC-MS IT-TOF. Extent of reaction was monitored by thin layer chromatography (TLC) using Merck 60 F_{254} pre-coated silica gel plates with fluorescent indicator UV254. After the plates were subjected to elution in the TLC chamber, the spots were visualized under UV light or using the appropriate stain (I₂, KMNO₄, ninhydrin, ceric ammonium molybdate (CAM)). Flash column chromatography was carried out using Merck silica gel (0.040-0.063).



2. Synthetic procedure and LC-MS profile of alkyne-functionalized inhibitor

Scheme S1 Synthesis of alkyne-functionalized inhibitor L1.

Procedure to (S)-tert-butyl 1-(1-methyl hydrazinyl)-1-oxo-3-phenylpropan-2-yl carbamate (2). N-Boc-L–phenylalanine (1) (1 eq.; 3 mmol; 0.80 g) was dissolved in dry THF (20 mL) and cooled to -25°C in an acetone/dry ice bath under nitrogen atmosphere. To the stirred solution, N-methylmorpholine (NMM) (1.2 eq.; 3.6 mmol; 0.40 mL), followed by isobutyl chloroformate (ISCF) (1.2 eq.; 3.6 mmol; 0.47 mL) was added dropwise to the reaction mixture. The reaction was stirred for 3h and filtered to remove the NMM.HCl salt. Methylhydrazine sulphate (1.2 eq.; 3.6 mmol; 0.52 g) was dissolved in H₂O (1 mL), and 5N NaOH (2 mL) was added under ice cooling. The methylhydrazine solution was added dropwise to the filtrate at -25°C. The reaction was allowed to warm to room temperature and stirred for 18h. After evaporation of the solvent, the aqueous residue was extracted with EA (3 x 50 mL). Organic phases were pooled, washed with saturated NaHCO₃ (2 X 30 mL), brine (2 X 30 mL), dried over Na₂SO₄ and concentrated in vacuo. Purification by flash column chromatography (EA:Hexane, 1:1) afforded (2) as a colourless oil (0.45g, 51% yield). ¹H NMR (300 MHz, CDCl₃) δ 1.41 (s, 9H), 2.50 (bs, 2H), 3.07 (s, 3H), 3.34 (s, 2H), 5.49 (m, 1H), 7.20 (bs, 1H), 7.23-7.29 (m, 5H). ¹³C NMR (75 MHz, CDCl₃) δ 28.3, 38.5, 40.0, 50.8, 79.3, 126.6, 128.1, 129.4, 155.2. LC-MS: m/z [M+H]⁺ calcd: 294.17, found: 294.16.

Procedure (S)-tert-butyl 1-(1,2-dimethyl hydrazinyl)-1-oxo-3-phenylpropan-2-yl to carbamate (4). To a solution of (2) (1 eq.; 0.58 mmol; 0.17 g) in dry THF (10 mL) was added formaldehyde (37 wt% in H_2O) (1.2 eq.; 0.70 mmol; 0.05 mL), followed by anhydrous MgSO₄ (1 eq.; 0.58 mmol; 0.07 g) and glacial acetic acid (0.3 eq.; 0.19 mmol; 0.02 mL). The reaction was stirred at room temperature for 4h. After removal of the solvent, resulting residue was dissolved in EA, washed with brine (3 x 50 mL), dried over Na₂SO₄ and concentrated in vacuo to give (S)-tertbutyl 1-(1-methyl-2-methylenehydrazinyl)-1-oxo-3-phenylpropan-2-ylcarbamate (3) a crude colourless oil and was used directly for next step without purification. (3) was dissolved in dry THF (10 mL) and cooled to 0°C under an ice water bath. Glacial acetic acid (6 mL), followed by NaBH₃CN (2.5 eq.; 1.45 mmol; 0.09 g) was added. The reaction was allowed to warm to room temperature and stirred overnight. After evaporation of solvent, resulting residue was cooled to 0°C and saturated NaHCO₃ was added. Aqueous solution was extracted with EA (3 x 50 mL). The combined EA layers were washed with NaHCO₃ (2 x 50 mL), brine (2 x 50 mL), dried over Na₂SO₄ and concentrated in vacuo. Residue was dissolved in MeOH (13 mL) and 1N NaOH (7 mL) was added. The reaction was stirred at room temperature for 2h and concentrated in vacuo. The aqueous residue was extracted with EA (3 x 50 mL). Organic phases were pooled, washed with brine (2 x 50 mL), dried over Na₂SO₄ and concentrated in vacuo. Purification by flash column chromatography (EA:Hexane, 1:1) afforded (4) as a white solid (0.12g, 68% yield over two steps). ¹H NMR (500 MHz, CDCl₃) δ 1.28 (s, 9H), 2.45 (d, J = 6.30 Hz, 3H), 2.79 (s, 1H), 2.97-3.01 (m, 2H), 3.04 (s, 3H), 5.25 (m, 1H), 5.41 (m, 1H), 7.20-7.29 (m, 5H). ¹³C NMR (125 MHz, CDCl₃) δ 28.3, 31.8, 35.4, 39.9, 50.9, 79.3, 126.5, 128.1, 129.5, 138.5, 174.1. LC-MS: m/z [M+H]⁺ calcd: 308.19, found: 308.18.

Procedure to (S)-2-amino-N,N'-dimethyl-3-phenylpropanehydrazide (5). A solution of **(4)** (1 eq.; 2.39 mmol; 0.73 g) in dry DCM (8 mL) was cooled to 0°C under an ice water bath and TFA (2 mL) was added dropwise. The reaction was allowed to warm to room temperature and stirred for 3h. After concentration in vacuo, resulting residue was neutralized to pH 10 using 1N NaOH. Aqueous solution was extracted with DCM (3 x 50 mL). The combined DCM layers were washed with brine (2 x 50 mL), dried over Na₂SO₄ and concentrated in vacuo, giving **(5)** as a colourless oil (0.46g, 93% yield). ¹H NMR (500 MHz, CDCl₃) δ 2.17 (bs, 2H), 2.56 (s, 3H), 2.76 (dd, J = 13.2, 8.8 Hz, 1H), 2.87 (m, 1H), 2.95 (s, 1H), 3.08 (s, 3H), 4.62 (m, 1H), 7.20-7.31 (m, 5H). ¹³C NMR

(125 MHz, CDCl₃) δ 32.2, 35.4, 40.9, 52.1, 126.9, 128.6, 129.4, 137.4, 176.7. LC-MS: m/z [M+H]⁺ calcd: 208.14, found: 208.13.

Procedure to (S)-N-(1-(1,2-dimethylhydrazinyl) -1-oxo-3-phenylpropan-2-yl)-4-(prop-2-ynyl oxy)benzamide (6). 4-(prop-2-ynyloxy)benzoic acid (1.2 eq.; 0.71 mmol; 0.10 g) was dissolved in dry DCM (10 mL) and cooled to 0°C under an ice water bath. To the stirred solution, HBTU (1.2 eq.; 0.71 mmol; 0.27 g), followed by DIEA (1.2 eq.; 0.71 mmol; 0.12 mL) was added to the reaction mixture. The reaction was stirred 10 min, after which **(5)** (1 eq.; 0.60 mmol; 0.12 g) was added. The reaction was stirred overnight at room temperature. After evaporation of the solvent, the residue was suspended in water and extracted with EA (3 x 50 mL). Organic phases were pooled, washed with saturated NaHCO₃ (2 X 50 mL), 1N HCI (2 x 50 mL), brine (2 x 50 mL), dried over Na₂SO₄ and concentrated in vacuo. Purification by flash column chromatography (EA:Hexane, 1:1) afforded **(6)** as a colourless oil (0.21g, 94% yield). ¹H NMR (300 MHz, DMSO-d₆) δ 2.57 (d, J = 5.4 Hz, 3H), 2.89-3.05 (m, 2H), 3.00 (s, 3H), 3.57 (m, 1H), 4.85 (d, J = 2.13 Hz, 2H), 5.49 (m, 1H), 7.02 (d, J = 8.55 Hz, 2H), 7.14-7.34 (m, 5H), 7.81 (d, J = 8.52 Hz, 2H), 8.33 (d, J = 8.55 Hz, 1H). ¹³C NMR (75 MHz, DMSO-d₆) δ 31.1, 35.0, 36.7, 51.6, 55.6, 78.5, 79.0, 114.3, 126.2, 127.2, 128.1, 129.1, 129.2, 139.1, 159.5, 165.7, 173.4. LC-MS: m/z [M+H]⁺ calcd: 366.17, found: 366.21.

Procedure to (S)-N-(1-(2-cyano-1,2-dimethyl hydrazinyl)-1-oxo-3-phenylpropan-2-yl)-4-(prop-2-ynyloxy)benzamide (L1). To a solution of (6) (1 eq.; 0.56 mmol; 0.21 g) in dry MeOH (10 mL) was added sodium acetate (2.8 eq.; 1.57 mmol; 0.13 g), followed by cyanogen bromide (3.3 eq.; 1.85 mmol; 0.20 g). The reaction was stirred overnight at room temperature. After evaporation of the solvent, residue was suspended in water and extracted with EA (3 x 50 mL). Organic phases were pooled, washed with brine (2 x 50 mL), dried over Na₂SO₄ and concentrated in vacuo. Purification by flash column chromatography (EA:Hexane, 1:1) afforded (L1) as a yellow oil (0.07g, 30% yield). ¹H NMR (300 MHz, CDCl₃) δ 2.53 (s, 1H), 3.05 (m, 2H), 3.23 (s, 3H), 3.33 (s, 3H), 4.72 (d, J = 2.3 Hz, 2H), 5.47 (m, 1H), 6.61 (d, J = 7.89 Hz, 1H), 6.96 (d, J = 8.88 Hz, 2H), 7.23-7.35 (m, 5H), 7.66 (d, J = 8.70 Hz, 1H), 7.76 (d, J = 8.52 Hz, 1H). ¹³C NMR (75 MHz, CDCl₃) δ 30.5, 38.1, 41.2, 50.6, 55.8, 76.0, 77.8, 114.7, 124.3, 126.7, 127.4, 128.8, 129.2, 129.6, 135.5, 160.3, 166.8, 173.3. LC-MS: m/z [M+H]⁺ calcd: 391.17, found: 391.16.

S4



3. Synthetic procedure and LC-MS profiles of azide-functionalized localization peptides

3 microreactors containing Rink amide resin (150 mg; loading = 0.6 mmol/g) were swelled in DMF for 2h. The solvent was removed, followed by Fmoc-deprotection with piperidine (20 wt% in DMF; 20 mL). The reaction was shaken at room temperature for 2h. The resulting resins were washed with DMF (3 x 50 mL), DCM (3 x 50 mL), MeOH (3 x 50 mL) and DMF (1 x 50 mL). The extent of reaction was monitored by ninhydrin test, where the presence of a primary amine is indicated by the blue colour of the resin. The resins were then added to pre-activated solutions of Fmocprotected amino acid (4 eq.; 0.36 mmol), HOBT (4 eq.; 0.36 mmol; 0.05 g), HBTU (4 eq.; 0.36 mmol; 0.14 g) and DIEA (4 eq.; 0.36 mmol; 0.12 mL) in DMF (15 mL). The reactions were shaken overnight at room temperature. The resulting resins were washed with DMF (3 x 50 mL), DCM (3 x 50 mL), MeOH (3 x 50 mL) and DMF (1 x 50 mL). The extent of reaction was monitored by ninhydrin test, where the absence of primary amine is indicated by the red colour of the resin. Repeated cycles of Fmoc-deprotection and coupling to Fmoc-protected amino acids resulted in peptide elongation. Finally, the N-terminus of each resin-bound peptide was coupled to 4azidobutanoic acid (4 eq.; 0.36 mmol; 0.05 g) in the presence of HOBT (4 eq.; 0.36 mmol; 0.05 g), HBTU (4 eq.; 0.36 mmol; 0.14 g) and DIEA (4 eq.; 0.36 mmol; 0.12 mL) in DMF (15 mL). After a round of washings, the resins were dried under vacuum. The dried resins were transferred from their microreactors to separate reaction vessels. All three peptides were cleaved and deprotected from the solid support using TFA:TIS:H₂O (95:2.5:2.5). Ether precipitation, followed by purification with preparatory HPLC, afforded the peptides as white solids. The peptides were characterized

using LC-MS IT-TOF. The m/z values obtained were fractions of the mass of target peptides due to positive charges present on the peptides, contributed by the protonation of arginine and lysine residues.

N₃-Tat: N₃-(CH₂)₃-CONH-RKKRRQRRR-CONH₂

LC-MS: m/z [M+H]⁺ calcd: 1449.92. [M+H]³⁺ calcd: 483.97, found: 483.97.

N₃-SV40: N₃-(CH₂)₃-CONH-KKKRKV-CONH₂

LC-MS: m/z [M+H]⁺ calcd: 896.62. [M+H]²⁺ calcd: 448.81, found: 448.80.

N₃-KK: N₃-(CH₂)₃-CONH-KK(palmitoyl)-CONH₂

LC-MS: $m/z [M+H]^+$ calcd: 623.49, found: 623.48.









4. Synthetic procedure and LC-MS profiles of "click" inhibitor-peptide conjugates

General procedure for "click" assembly of inhibitor-peptide conjugates. To a 48 μ L solution of DMSO:H₂O (1:1) / ^tBuOH:H₂O (1:1) in a 2 mL eppendorf tube, were added 8 μ L of alkyne-functionalized inhibitor L1 (20 mM) and 4 μ L of azide-functionalized localization peptide (20 mM). 8 μ L of CuSO₄.5H₂O (4 mM) and 8 μ L of TBTA (10 mM) were pre-mixed and added to the solution, followed by 4 μ L of sodium ascorbate (50 mM). The tubes were capped and shaken at room temperature for one day, after which products were purified by preparatory HPLC. The conjugates were re-dissolved in 80 μ L of 5% DMSO and analyzed by LC-MS IT-TOF. L1-Tat: LC-MS: m/z [M+H]⁺ calcd: 1840.09. [M+H]³⁺ calcd: 614.03, found: 614.37. L1-SV40: LC-MS: m/z [M+H]⁺ calcd: 1286.79. [M+H]²⁺ calcd: 643.90, found: 643.91. L1-KK: LC-MS: m/z [M+H]⁺ calcd: 1013.66, found: 1013.67.



L1-Tat







L1-Tat $\begin{array}{c} & & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & & \\ & & & \\ & & &$

L1-SV40



L1-KK





5. Procedure for IC₅₀ evaluation against recombinant cathepsin L

 IC_{50} measurements were performed to determine the potency of the alkyne-functionalized inhibitor **L1** and its three conjugates against recombinant cathepsin L. The dose-dependent inhibition assays were performed by varying the concentration of inhibitors under fixed enzyme and fluorogenic substrate Z-FR-AMC (Biotium) concentrations of 5 nM and 20 µM respectively. The fluorometric assays were performed in 384-well Greiner black plates, and the fluorescence readouts were monitored over 3h assay runs using BioTek Multi-Mode Microplate Reader for fluorescence background at an excitation wavelength of 360 nm and an emission wavelength of 460 nm. Kinetic endpoints were obtained and analyzed using GraphPad Prism software to plot out the IC₅₀ graphs (Figure 2a).

6. Cell culture

The HepG2 cell line was cultured in growth media (DMEM) supplemented with 10% fetal bovine serum and 1% penicillin. Cells were maintained in a humidified atmosphere of 5% CO₂ at 37°C.

7. Procedure and results for western blot of HepG2 cell lysate



Figure S2 Western blot results of HepG2 cell lysate with anti-cathepsin L. Lane 1: 50 ng recombinant cathepsin L (Heavy chain – 24 kD). Lane 2: 10 µg HepG2 cell lysate (Heavy chain – 21 kD).

An SDS-PAGE gel of HepG2 cell lysate was electrotransferred overnight at 4 °C to a PVDF membrane. The membrane was blocked with 3% BSA in 0.1% Tween 20 (TBS-T) for 1 h at 25 °C. After four 10 min washes with TBS-T, the membrane was treated with cathepsin L antibody (Abcam) in 3% BSA in TBS-T (1:600 dilution) for 1 h at 25 °C. The membrane was then washed four times (10 min each wash) with TBS-T and subsequently incubated with goat anti-mouse HRP antibody (Pierce) (1:3000 dilution) for 1 h at 25 °C. The membrane was washed four times (10 min each wash) with TBS-T. Enhanced ChemiLuminescent (ECL) reagent was used and the image was acquired using dark room development facilities.

8. Procedure for IC₅₀ evaluation against HepG2 cell lysate

 IC_{50} measurements were performed to determine the potency of the alkyne-functionalized inhibitor **L1** and its three conjugates against HepG2 cell lysate. The dose-dependent inhibition assays were performed by varying the concentration of inhibitors under fixed lysate and fluorogenic substrate Z-FR-AMC (Biotium) concentrations of 0.1 mg/mL and 5µM respectively. The fluorometric assays were performed in 384-well Greiner black plates, and the fluorescence readouts were monitored over 3h assay runs using BioTek Multi-Mode Microplate Reader for fluorescence background at an excitation wavelength of 360 nm and an emission wavelength of 460 nm. Kinetic endpoints were obtained and analyzed using GraphPad Prism software to plot out the IC₅₀ graphs (Figure 2b).

9. Procedure for cell imaging with control peptides





HepG2 cells were seeded on glass-bottom dishes (Mattek) and grown till 70-80% confluency. The growth medium was removed; the cells were first fixed with 3.7% paraformaldehyde in PBS for 15 min and washed twice with PBS. Next the fixed cells were incubated with 1 µM SG-peptides for 2h. The cells were then washed twice with PBS and incubated with 0.25 µg/mL of organelle trackers according to the various SG-peptides: Lyso Tracker Red (Invitrogen) for **SG-Tat**; Hoechst stain (Cell Technology) for **SG-SV40**; Cell mask orange (Invitrogen) for **SG-KK**. After washing the cells with PBS, the cells were imaged with an Olympus IX71 inverted microscope, equipped with a 60X oil objective (NA 1.4, WD 0.13 mm) and CoolSNAP HQ CCD camera (Roper Scientific). Images were processed with MetaMorph software (Molecular Devices) (Figure 3a).



Figure S4 Alternative procedure: Live HepG2 cells were first treated with 1 μ M **SG-Tat** for 2h, followed by incubation with 0.25 μ g/mL of Lyso Tracker Red. Cells were then fixed with 3.7% paraformaldehyde in PBS for 15 min and imaged with Olympus BX61 confocal microscope. Scale bar = 20 μ m.

10. Procedure and results for cell imaging with cell-permeable fluorogenic substrate



Figure S5 Fluorescence images of fixed HepG2 cells after incubation with 2% DMSO, followed by cell-permeable cathepsin L substrate, and finally Lyso Tracker Green and Hoechst stain.

HepG2 cells were seeded on glass-bottom dishes (Mattek) and grown till 70-80% confluency. The growth medium was removed; the cells were fixed with 3.7% paraformaldehyde in PBS for 15 min and washed twice with PBS. Next the cells were incubated with 0.1 μ M of alkyne-functionalized inhibitor **L1** and peptide conjugates for 2h. The cells were then washed twice with PBS and incubated with 2.6 μ M of cell-permeable cathepsin L substrate (Cell Technology) for 30 min. After the cells were washed twice with PBS, they were incubated with 0.25 μ g/mL of Lyso Tracker

Green (Invitrogen), and subsequently 0.25 µg/mL of Hoechst stain (Cell Technology). The cells were washed with PBS and imaged with an Olympus IX71 inverted microscope, equipped with a 60X oil objective (NA 1.4, WD 0.13 mm) and CoolSNAP HQ CCD camera (Roper Scientific). Images were processed with MetaMorph software (Molecular Devices) (Figure 3b).

11. Procedure and results for flow cytometry analysis

HepG2 cells were harvested, washed once with cold PBS, and fixed with 3.7% paraformaldehyde in PBS for 15 min. Cells were washed once with cold PBS and resuspended in sodium acetate buffer (25 mM sodium acetate, 1.25 mM EDTA, pH 5.5). Cells were then incubated with 1 μ M alkyne-functionalized inhibitor **L1** and peptide conjugates for 2h, followed by 20 μ M of cell-permeable cathepsin L substrate (Cell Technology) for 2h. Flow cytometry analysis was performed on Guava EasyCyte Plus machine equipped with a 532 nm excitation laser and a 690/50 nm detection filter. Histograms were generated using Guava Express Pro analysis software. Mean fluorescence intensity (MFI) was computed based on events marked red as "M1" in plots shown in Figure S6.



Figure S6 Flow cytometry histograms after incubation with 1 μ M of the indicated inhibitors, followed by addition of cell-permeable fluorogenic substrate.

Inhibitor	MFI	Relative MFI
No inhibitor	491.89	1.00
L1	394.31	0.80
L1-Tat	362.60	0.74
L1-SV40	424.53	0.86
L1-KK	382.13	0.78

Table 1 Mean fluorescence intensity (MFI) from flow cytometry analysis.

12. ¹H and ¹³C NMR spectra

1H normal range AC 300 Compound 2





ī.

I.



S15



ī.





1H normal range AC 300 Compound 6

ī.





*** Current Data Parameters *** NAME : se24su~1 EXPN 0 : 9 PR O C N O: 1 *** Acquisition Parameters *** IN STRUM: spect LOCNUC: 2 H N S : 8 N U C LEU S o ff 01 : 1853.43 Hz PULPROG zg30 SFO1 3:00.1318534 MHz SOLVEN T: D M SO SW : 17.9519 ppm TD : 32768 *** Processing Parameters *** LB : 0.30 H z *** 1D NMR Plot Parameters * NUCLEUS off



*** Current Data Parameters *** NAME : se24su~1 EXPN 0 : 10 PROCNO: 1 *** Acquisition Parameters *** IN STR UM: spect LOCNUC: 2 H NS : 49 N U C LEU S o ff 01 : 7924.11 Hz PULPROG zgpg30 SFO1 :75.4756731 MHz SOLVEN T: D M SO SW : 238.2968 ppm TD : 32768 *** Processing Parameters *** LB : 1.00 H z *** 1 D N M R Plot Parameters ** NUCLEUS off

Bruker



S18